



**QUEEN'S
UNIVERSITY
BELFAST**

Gender Biases in (Inter) Action: The Role of Interviewers' and Applicants' Implicit and Explicit Stereotypes in Predicting Women's Job Interview Outcomes

Latu, I. M., Schmid Mast, M., & Stewart, T. L. (2015). Gender Biases in (Inter) Action: The Role of Interviewers' and Applicants' Implicit and Explicit Stereotypes in Predicting Women's Job Interview Outcomes. *Psychology of Women Quarterly*, 39(4), 539-552. <https://doi.org/10.1177/0361684315577383>

Published in:
Psychology of Women Quarterly

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights
© 2015 The Authors.
This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Gender Biases in (Inter)Action: The Role of Interviewers' and Applicants' Implicit and
Explicit Stereotypes in Predicting Women's Job Interview Outcomes

Ioana M. Latu

Rutgers University

Marianne Schmid Mast

University of Lausanne, Switzerland

Tracie L. Stewart

Kennesaw State University

Author Note

Ioana M. Latu, Department of Psychology, Rutgers University; Marianne Schmid Mast, Department of Organizational Behavior, University of Lausanne, Switzerland; Tracie L. Stewart, Department of Psychology, Kennesaw State University.

The authors would like to thank the students of the "Workplace Interactions" Seminar from University of Neuchâtel, Switzerland for data collection and Camille Morgenthaller for coding. We also thank Dario Bombari, Denise Frauendorfer, Sebastian Otte, and Michael Triplett for feedback on the manuscript.

Correspondence concerning this manuscript should be addressed to Ioana Latu, Rutgers University, 311 N. Fifth Street, Camden, NJ 08102. Email: Ioana.Latu@rutgers.edu

Abstract

Although explicit stereotypes of women in the workplace have become increasingly positive (Duehr & Bono, 2006), negative stereotypes persist at an implicit level, with women being more likely associated with incompetent—and men with competent—managerial traits (Latu et al., 2011). Drawing upon work on self-fulfilling prophecies and interracial interactions, we investigated whether and how implicit and explicit gender stereotypes held by both male interviewers and female applicants predicted women's interview outcomes. Thirty male interviewers conducted mock job interviews with 30 female applicants. Before the interview, we measured interviewers' and applicants' implicit and explicit gender stereotypes. The interviewers' and applicants' implicit stereotypes independently predicted external evaluations of the performance of female applicants. Whereas female applicants' higher implicit stereotypes *directly* predicted lower performance, male interviewers' implicit stereotypes *indirectly* impaired female applicants' performance through lower evaluations by the interviewer and lower self-evaluations by the applicant. Moreover, having an interviewer who was at the same time high in implicit and low in explicit stereotypes predicted the lowest performance of female applicants. Our findings highlight the importance of taking into account both implicit and explicit gender stereotypes in mixed-gender interactions and point to ways to reduce the negative effects of gender stereotypes in job interviews.

Keywords: stereotyped attitudes, sex role attitudes, employment discrimination, job performance, self-fulfilling prophecies

Gender Biases in (Inter)Action: The Role of Interviewers' and Applicants' Implicit and Explicit Stereotypes in Predicting Women's Job Interview Outcomes

Considerable progress toward achieving gender equality in the United States and Western Europe has been attained in the last few decades. The number of women in leadership positions has been slowly rising. For example, women held 16.9% of U.S. Fortune 500 companies' board seats in 2013, compared to 9.6% in 1995 (Catalyst, 2011). Women also lead several powerful European nations such as Germany, Switzerland, Denmark and Norway. Moreover, organizations are increasingly committed to promoting gender equality through diversity programs and policies. However, despite an improvement in explicit gender stereotypes (Duehr & Bono, 2006; Stoker, Van der Velde, & Lammers, 2012), negative stereotypes persist at an implicit level, with women being more likely to be implicitly associated with negative managerial traits (Latu et al., 2011) and with roles incompatible with leadership (Nosek, Banaji, & Greenwald, 2002; Rudman & Kilianski, 2000).

One pivotal point at which gender stereotypes can manifest themselves in workplace inequality is in the door to employment itself: the job interview. But what is the role that gender stereotypes play in female job applicants' interview performance and employers' interview-based job decisions? More specifically, if stereotyping contributes to gender inequality in such decisions, is it the interviewers' or female applicants' stereotypes of women in the workplace that drive the effects, and is it implicit or explicit stereotypes that have the most influence? Drawing upon prior empirical work on bias in interracial interactions (Dovidio, 2001; Dovidio, Kawakami, & Gaertner, 2002) as well as self-fulfilling prophecies (Rosenthal & Jacobson, 1968; Word, Zanna, & Cooper, 1974), we designed a paradigm to assess the effects of gender stereotypes held by both male interviewers and female applicants on female applicants' interview performance in a simulated job interview.

Consistent with prior work on interracial interactions that investigates how Whites' racial biases affect their behavior towards Blacks (Dovidio, 2001; Dovidio et al., 2002; Penner et al., 2010), we focused exclusively on mixed-gender dyads in which interviewers were always men and applicants were always women. In the case of gender biases, our decision is strengthened by the fact that, in previous research, it was men rather than women who held negative implicit (Latu et al., 2011) and explicit gender-managerial stereotypes (Duehr & Bono, 2006; Schein, Mueller, & Jacobson, 1989).¹

In the present study, we sought to address three main questions. First, how do interviewers' implicit stereotypes influence the performance of female applicants in job interviews? Second, do interviewers' implicit and explicit stereotypes interact to predict performance? Finally, do applicants' own stereotypes predict their performance in job interviews?

Implicit Gender Stereotypes

Whereas past research showed that female managers were assigned more negative attributes compared to male managers (Heilman, Block, & Martell, 1995) and successful managers were perceived as possessing traits more commonly ascribed to men rather than women (e.g., “think manager—think male paradigm”; Schein et al., 1989; Schein, Mueller, Lituchy, & Liu, 1996), recent research has shown increasingly favorable views of women in the workplace (Koenig, Eagly, Mitchell, & Ristikari, 2011; Stoker et al., 2012). For example, male managers perceived women as more assertive, more ambitious, and less submissive than in the past (Duehr & Bono, 2006).

These increasingly equitable views of gender roles in western cultures may reflect an actual change in people's thoughts about women and leaders. However, given the current societal norms that increasingly condemn stereotyping and prejudice, people are likely to deny having personal biases and control the extent to which they express those biases. Thus,

the improvement in self-reported gender stereotypes may not reflect an absolute change, but rather social desirability concerns because people are less willing to declare their negative thoughts about women and deliberately control their sexist responses.

To bypass such social desirability concerns, researchers have developed implicit measures of gender biases. These measures assess people's relatively automatic mental associations between women and certain traits. Within the workplace and leadership context, recent research has shown that people tend to hold implicit mental associations that can put women with leadership aspirations at a disadvantage. For example, men are more likely to be associated with work, authority roles, hierarchies, power, and agentic traits, whereas women are linked with family, subordinate roles, egalitarian structures, warmth, and communal traits (Nosek et al., 2002; Rudman & Glick, 2001; Rudman & Kilianski, 2000; Rudman & Phelan, 2010; Schmid Mast, 2004). Not only are women less likely than men to be associated with managerial roles and domains, but also, once in this domain, they are implicitly viewed in a negative light, especially by male respondents. For example, Latu and colleagues (2011) found that men were more likely to associate men with successful manager traits (e.g., leader, competent, knowledgeable) and women with unsuccessful manager traits (e.g., follower, incompetent, ignorant), compared to reversed pairings.

How harmful are these implicit associations for women who aspire to become leaders? Although we know that implicit stereotypes can predict the decisions and discriminatory behaviors of the person holding these biases (Latu et al., 2011; Rudman & Glick, 2001; Williams, Paluck, & Spencer-Rodgers, 2010), we do not know how implicit stereotypes play out in actual social interactions between men and women. In particular, we do not know how implicit and explicit gender biases influence not only the decisions of the individuals holding these biases but also the behavior of the women who are the targets of those biases. Second, in the context of an actual social interaction between men and women

such as the job interview, we do not know which of the dyad members' views of women in the workplace best predict behavioral outcomes: the male interviewers', the female applicants', or both?

Gender Stereotypes and Self-Fulfilling Prophecies

There is reason to believe that interviewers' gender stereotypes can, indeed, predict the performance of female job applicants. In fact, research on self-fulfilling prophecies in job interviews suggests that interviewers' expectations about applicants are expressed in subtle behavioral cues, which then bring about the expected behavior (Rosenthal & Jacobson, 1968).² For example applicants interacting with interviewers displaying unfriendly nonverbal behaviors performed less well compared to those interacting with interviewers displaying friendly nonverbal behaviors (Word et al., 1974). Within the realm of mixed-gender interactions, female applicants behaved more flirtatiously when interviewed by a male interviewer who was led to believe that she was attracted to him compared to dyads in which this expectation was not induced (Ridge & Reber, 2002). Finally, applicants' expectations of the interviewer also affected the applicants' behavior such that female applicants who believed they were interviewed by a male interviewer with traditional gender beliefs spoke less during the interview, dressed more feminine, and were rated as more attractive compared to applicants who believed they were interviewed by a nontraditional man (von Baeyer, Sherk, & Zanna, 1981).

Overall, research on self-fulfilling prophecies shows that interviewers' expectations, communicated through certain nonverbal behaviors, can influence how applicants perform. Based on this theoretical framework, we proposed that male interviewers' gender stereotypes (which guide their expectations) would, indeed, predict the performance of female job applicants, similar to a self-fulfilling prophecy. But is this self-fulfilling prophecy driven by

interviewers' implicit or explicit gender stereotypes? In the present study we investigated both the independent and interactive effects of interviewers' implicit and explicit stereotypes.

Hypothesis 1: Interviewers' Implicit Gender Stereotypes

We propose that when it comes to independent effects, implicit gender stereotypes held by interviewers would be more likely than explicit stereotypes to predict applicants' performance, given that implicit stereotypes are less likely to be influenced by social desirability concerns. But what is the process involved? So far, we know that implicit gender stereotypes can predict the decisions and discriminatory behaviors of the person holding these biases (Latu et al., 2011; Rudman & Glick, 2001; Williams et al., 2010). For example, the more participants associated men, more than women, with managerial competence, the more they favored male employees in a hypothetical salary allocation scenario (Latu et al., 2011). We similarly propose that in the context of a job interview, interviewers' implicit gender stereotypes would predict how interviewers evaluate female applicants, such that more negative implicit stereotypes would predict lower interviewer evaluations.

We extended this chain of reasoning and hypothesized that applicants would pick up on their interviewers' evaluations and assimilate them, such that lower interviewer evaluations would predict lower applicant self-evaluations. We hypothesized that applicants' self-evaluations would be influenced by interviewers' evaluations because applicants in a stressful, evaluative situation are likely to be sensitive to cues coming from interviewers—cues that may contain feedback about their performance.

In turn, given that self-evaluations of performance tend to predict how people are evaluated by others (John & Robins, 1994), we hypothesized that applicants' self-evaluations would predict how they are evaluated by external observers, thus leading to a full circle of self-fulfilling prophecies. In other words, we proposed that interviewers' implicit gender stereotypes would not necessarily predict external performance evaluations of female job

applicants directly, but rather largely indirectly through interviewers' evaluations and applicants' self-evaluations. More specifically, the greater interviewers' tendency to associate women, more than men, with incompetence, the less hireable they will evaluate the female applicant. In turn, the interviewers' evaluations will predict how the applicants evaluate their own performance, such that the more negatively the interviewers evaluate the applicants, the more negatively applicants will evaluate themselves. Finally, these negative self-evaluations will predict more negative impressions from external raters, such that applicants will be rated as less hireable by external raters, thus leading to a self-fulfilling prophecy. Thus, Hypothesis 1 predicted that male interviewers' implicit stereotypes will indirectly predict external performance evaluations of female applicants, such that the effect of interviewers' implicit stereotypes on applicant's performance (as evaluated by naïve, external raters) will be mediated, in a serial fashion, by interviewers' evaluations of the applicant and applicants' self-evaluations.

Hypothesis 2: Interviewers' Implicit and Explicit Gender Stereotypes

We also hypothesized that the effects of implicit gender stereotypes on applicants' performance would be qualified by an interaction with explicit stereotypes. Within interracial interactions, there is evidence of interactive effects of Whites' implicit and explicit racial biases on Black targets' perceptions and behaviors. Given that explicit racial biases tend to be expressed through verbal behaviors, whereas implicit racial biases tend to be expressed through nonverbal behaviors (Dovidio et al., 2002), Whites, who explicitly hold positive attitudes toward Blacks but implicitly hold negative attitudes, may display incongruent verbal and nonverbal behaviors. This group was called *aversive racists* by Dovidio and his colleagues (Dovidio & Gaertner, 2000; Penner, et al., 2010). Research showed that aversive racism has the most negative consequences in interracial interactions. For example, White/Black dyadic teams showed the worst problem-solving performance when White

participants were high in implicit and low in explicit biases (Dovidio, 2001). Similarly, Black patients had the most negative impressions of a medical interaction when their non-Black physicians were high in implicit and low in explicit racial biases (Penner et al., 2010). These findings may be surprising given that we would expect a combination of high implicit and high explicit biases to predict negative outcomes. However, what makes interactions with aversive racists difficult is that they send mixed, incongruent signals. Their overt (verbal) behaviors, which are an expression of their positive explicit biases, tend to be friendly. Conversely, their subtle, spontaneous behaviors, which are an expression of their negative implicit biases, tend to be unfriendly. This inconsistency may be perceived as deceitful and may interfere with the smooth functioning of the social interaction, thus resulting in decreased performance and negative impressions from the target of this bias.

We similarly propose that in mixed-gender interactions, the performance of female job applicants would be most negatively affected when interacting with male interviewers who are high in implicit gender biases but low in explicit gender biases. These *aversive sexist* interviewers may display positive signals verbally and negative signals nonverbally. Resolving this inconsistency may be taxing for female applicants, thus reducing their performance. For applicants' own implicit and explicit gender stereotypes, there would be no interaction in predicting applicants' own performance, given that mixed verbal and nonverbal signals only have an effect when coming from an interaction partner (i.e., the interviewer) in an actual social interaction. Thus Hypothesis 2 predicts that interviewers' implicit and explicit gender stereotypes will interact to predict female applicants' job interview performance, such that female applicants will show the lowest performance when interacting with male interviewers who hold high implicit and low explicit gender stereotypes.

Hypothesis 3: Applicants' Implicit Gender Stereotypes

It might also be the case that applicants' own gender stereotypes can influence how they perform. Although women show less negative stereotypes of women in workplace settings, both implicitly (Latu et al., 2011) and explicitly (Schein et al., 1989), the variance in women's own stereotypes can still predict performance outcomes.

The idea that women's implicit self-stereotypes can predict women's actual performance on a relevant task is supported by work showing that the implicit activation of negative self-stereotypes can affect performance. For gender stereotypes, the implicit activation of Asian women's gender identity negatively affected their math performance, whereas the implicit activation of their Asian identity increased their math performance (Shih, Pittinsky, & Ambady, 1999). However, priming stereotypes causes participants to be aware of stereotypes rather than to endorse them. There are, to date, no known studies showing that endorsing negative stereotypes about their gender group would predict women's performance on relevant tasks. We aim to eliminate this gap by proposing that women's negative implicit gender stereotypes would predict decreases in women's performance in a job interview.

Whereas for interviewers' implicit gender stereotypes we proposed an indirect effect (Hypothesis 1), for applicants' implicit gender stereotypes we propose a direct effect on externally evaluated performance. In other words, the way women think about women may directly influence how they present themselves in an evaluative situation such as the job interview. Women who implicitly think that women do not make good managers may behave in a disempowered way, leading to their performance being more negatively evaluated by others. We do not have a reason to expect an independent effect of applicants' explicit stereotypes on female applicants' job interview performance, given that explicit measures of bias are likely to be influenced by social desirability concerns. Therefore, Hypothesis 3 predicts that female applicants' own implicit gender stereotypes will directly predict their

externally rated performance, such that the more female job applicants associate women with incompetence, the less competently they will perform in the job interview.

In summary, in the present study we investigated how implicit and explicit gender biases play out in mixed-gender job interviews. One key element that sets apart our research is that we investigated gender biases in an actual social interaction, and as such, we were able to assess simultaneously how both male interviewers' and female applicants' gender biases predicted applicants' performance. Moreover, we investigated how implicit and explicit stereotypes may interact in predicting female applicants' performance.

Method

Participants

Participants ($N = 76$) were from a Swiss University and the broader community outside the university. They participated in two ostensibly unrelated studies: the first in which we measured their implicit and explicit stereotypes and the second in which they participated in mixed-gender dyads in a mock job interview. Male participants were always assigned the interviewer role, and female participants were always assigned the applicant role. The interviewer and applicant did not know each other prior to the interview. Participants did not receive any payment or rewards for their participation.

Participants were not aware of their role assignments until the interview task began. Given that our main focus was on impressions of hireability and that these subjective ratings are highly dependent on language proficiency, we eliminated dyads in which either or both members were non-native French speakers (total of 8 dyads). This resulted in a total of 60 participants in 30 dyads.

Procedure

The experimenter informed participants they would take part in two unrelated studies, which were conducted together due to time constraints. Participants were told that the first

study consisted of computer-based tasks and questionnaires and that the second investigated how people behave in job interviews. As part of the first study, participants first completed a measure of implicit gender stereotypes, then a measure of explicit gender stereotypes, and lastly a measure of sexism (The Ambivalent Sexism Inventory, Glick & Fiske, 2001).³ After the sexism measurement phase and just prior to the job interview, the experimenter assigned participants to either the interviewer or the applicant role. The seemingly random assignment was arranged such that male participants were always assigned the interviewer role and female participants were always assigned the applicant role. Several steps were taken in order to minimize the possibility that participants connected the two studies: We told participants the studies were originally planned and conducted by two different researchers in the lab and ensured that the two “studies” took place at different tables in a large laboratory room.

During a 10-min preparation phase, in which the interviewer and applicant worked separately, both received the job description of a regional marketing manager for a large convenience store. The interviewers also received a list of questions that they could use during the interview if they wished (e.g., “What is your greatest strength?”; “Can you describe a professional experience which was particularly challenging for you?”). The applicant did not receive question suggestions, rather only the instruction that they should prepare a general strategy of presenting themselves for the proposed job. After the preparation phase, the experimenter guided participants to a big table where they sat across from each other for the duration of the job interview. They were videotaped using two cameras installed on the table, which allowed us to obtain separate videotapes of the interviewer and applicant. Following a short post-interview questionnaire measuring performance, the experimenter thanked, debriefed, and dismissed participants.

Measures

Implicit gender stereotypes. We used a sequential priming task adapted from the Successful Manager Implicit Association Test (Latu et al., 2011) to measure implicit associations between gender and competent (leader, competent, knowledgeable, consistent, self-confident, trustworthy, self-controlled, well-informed, intelligent, fair, purposeful, and skilled) and incompetent (follower, incompetent, ignorant, inconsistent, insecure, dishonest, reckless, uninformed, dense, biased, aimless, and unskilled) manager traits. This computer task was composed of 160 trials divided in two blocks, modeled after Dovidio and colleagues (2002). Each trial started with a fixation point in the middle of the screen, followed by a prime denoting either a male or female name, or the name of a type of building (e.g., house, cabin). The prime remained on the screen for 250ms, after which it was replaced by a trait—either a positive or a negative manager trait or a positive or negative building trait (e.g., spacious, leaky). Participants' task was to press one of two keyboard keys to indicate, as quickly and accurately as possible, if the trait described a person or a building. Response times were measured in milliseconds.

For analyses, we computed mean response times for categorizing competent and incompetent manager traits after being primed with male and female names. We eliminated response times that were three standard deviations above and below the mean and log-transformed the remaining response times in order to normalize the distribution of values. The final score was computed by subtracting the average response time to stereotype-consistent trials (male-competent; female-incompetent) from those to stereotype-inconsistent trials (male-incompetent; female-competent). If participants more strongly associate men with competence and women with incompetence (stereotype-consistent) compared to men with incompetence and women with competence (stereotype-inconsistent), they should be faster to respond to the consistent compared to the inconsistent trials. Thus, higher numbers

denote stronger men-competence, women-incompetence associations. Both the male interviewer and the female applicant completed this measure.

Explicit gender stereotypes. To measure the extent to which participants explicitly associate men and women with leadership, we asked participants to estimate the percentage of men and women in the business domain who are leaders. This is an abbreviated version of the explicit gender-stereotyping measure used in Latu and colleagues (2011), which included estimates for several other traits (e.g., competent, knowledgeable). In order to obtain a score of their explicit gender-leadership associations, we computed a difference score by subtracting the score for women from the score for men. Higher (positive) numbers denote stronger explicit associations between men and leadership, whereas lower (negative) numbers denote stronger explicit associations between women and leadership. Both the male interviewer and the female applicant completed this measure.

Interviewers' evaluation of applicants' performance. We operationalized performance as the likelihood that the applicant would be hired based on her behavior during the job interview, and we assessed it from three points of view: interviewer, applicant, and external coders. Interviewers' evaluations of applicants are commonly used in organizational settings for hiring and promotion decisions; moreover, meta-analytical studies show that they are significantly correlated with objective measures of performance—defined as direct measures of countable behaviors (overall correlation: $r = .39$, which increases to $.71$ when the measures assessed precisely the same performance dimension; Bommer, Johnson, Rich, Podsakoff, & MacKenzie, 1995). In our study, immediately after the interview, interviewers completed five items: “In my opinion, the applicant is qualified for this job”; “Based on our interaction, the applicant is competent for this job”; “I believe that the applicant would be able to complete all the duties of the job”; “The applicant performed well during this job interview”; and “Based on this job interview, I believe the applicant should be hired for the

job,” rated on a 7-point scale from 1 (*strongly disagree*) to 7 (*strongly agree*). A sixth item, “Overall, how would you rate the applicant’s strength as an applicant during the job interview?,” rated on a 7-point scale from 1 (*not at all strong*) to 7 (*extremely strong*), was also included. Responses across the six items were averaged into a final score (Cronbach’s $\alpha = .95$), with higher numbers denoting more positive impressions of the applicant.

Applicants’ self-evaluation. Applicants’ self-evaluations are not usually the basis for decisions in job interview outcomes; however, these self-evaluations seem to be related to objective performance measures in cognitive domains (showing a medium-size, significant correlation of .33; Freund & Kasten, 2012). Moreover, they provide incremental validity when used in conjunction with other measures in work-related domains (Prediger, 1999). In our study, immediately after the interview, the applicants completed four items: “I was not very successful in my job interview” (reversed scored); “I gave a competent impression during this job interview”; “In my opinion, I performed well during this job interview”; and “I believe that based on my interview, I should be hired for the job,” rated on a 7-point scale from 1 (*strongly disagree*) to 7 (*strongly agree*). A fifth item, “Overall, how would you rate your strength as an applicant during the interview?,” rated on a 7-point scale from 1 (*not at all strong*) to 7 (*extremely strong*), was also included. Responses across the five items were averaged to obtain a final score (Cronbach’s $\alpha = .91$), with higher numbers denoting higher self-evaluated performance.

External coder evaluations. External coders’ evaluations are very similar to a hiring decision. By having a third party also evaluate applicants’ performance, we took the evaluation outside the interviewer-applicant dyad. However, by using both interviewers’ and external coders’ evaluations, we sought to disentangle the direct effects of interviewers’ gender bias on their own evaluations of applicants (interviewers with more negative stereotypes may evaluate female applicants more negatively) and the dynamic effect of

interviewers' stereotypes on applicants performance in an actual social interaction (interviewers with more negative stereotypes communicate negative expectations to female applicants, which in turn decrease their externally evaluated performance).

In our study, applicant performance was coded by two female coders who were unaware of our study's hypotheses. Both coders watched videos of the applicants in which the interviewer was not visible and rated the applicants' performance using a global impression coding scale, from 1 (*It would be doubtful that someone would hire the applicant following this interview*) to 5 (*If someone watched this interview, they would hire this applicant without hesitation*). Both coders were instructed to pay attention to specific behaviors (smiling, nodding, looking straight ahead at the interviewer sitting directly across the table), which were previously identified as being significantly related to higher competence ratings in job interviews (Howard & Ferris, 1996). The first coder watched a thin slice of the applicants' videotaped interview (4-min slice from 00:30 to 04:30) and coded applicant performance based on video only ($M = 3.00$, $SD = 1.20$). Coding reliability was assessed with another coder for a subset of 10 videos, $r = .82$. The second coder watched the videotaped interview entirely (average interview duration in seconds $M = 616.62$, $SD = 241.60$) and coded performance based on both sound and video ($M = 3.78$, $SD = .94$). This coder also considered the quality of the verbal content, including persuasiveness and fluency. Coding reliability was assessed with another coder for a subset of eight videos, $r = .95$. These two coding strategies were employed because we were not sure whether the verbal content of the interview (which is highly influenced by the type of questions the interviewer asked) would skew performance evaluations. However, the two ratings were significantly correlated ($r = .58$, $p = .001$) and thus were averaged into a single score, with higher numbers denoting more positive performance evaluations.

Personal variables. We also measured three personal variables, which we thought were relevant for the job interview situation. First, we measured both interviewers' and applicants' age, which they self-reported at the end of the experience. Second, before the interview, we asked both dyad members to report the number of actual interviews in which they participated as applicants and the number of actual interviews in which they participated as interviewers. Third, given that our mock job interview was for a marketing position, we asked participants to self-report their marketing experience ("I have experience in the marketing domain") on a scale from 1 (*strongly disagree*) to 7 (*strongly agree*).

Results

Preliminary Analyses

We first investigated personal characteristics of dyad members in order to establish whether there were significant differences between interviewers and applicants other than gender. We initially conducted a MANOVA, which showed no overall effect, $F(4, 55) = 1.40$, $p = .25$. However, a follow-up ANOVA revealed a significant difference in age. Interviewers ($M = 25.8$, $SD = 3.91$) were significantly older than applicants ($M = 23.8$, $SD = 2.71$), $F(1, 58) = 5.31$, $p = .025$. Therefore, in all subsequent analyses we controlled for age. In terms of their experience with job interviews, men participated, on average, in 1.5 ($SD = 5.08$) interviews as interviewers, whereas women averaged .27 ($SD = 1.05$), but this difference was not significant, $F(1, 58) = 1.69$, $p = .20$. In addition, there was not a significant difference between dyad members in terms of their experience as applicants in job interviews, $F(1, 58) = 0.79$, $p = .38$: men participated, on average, in 4.27 interviews as applicants ($SD = 4.66$) and women in 3.33 interviews ($SD = 3.40$). Finally, there was no difference in marketing experience between men ($M = 2.03$, $SD = 1.35$) and women ($M = 1.73$, $SD = 1.44$), $F(1, 58) = 0.69$, $p = .41$. In Table 1 we present means, standard deviations, and correlations between all study variables, separately for male interviewers and female applicants.

Analysis Plan

Given the dyadic nature of our data, we used Kenny's Actor-Partner Interdependence Model (Cook & Kenny, 2005) that is designed to analyze data in a dyadic design. In the context of our design, this model allowed us to investigate whether interviewers' and applicants' implicit and explicit gender stereotypes independently and/or interactively predicted female applicants' performance. As a preliminary step, we investigated whether the interviewer and applicant scores were independent for each predictor (implicit gender stereotypes and explicit gender stereotypes). Given that the members of the dyads were distinguishable (each dyad member had a role that was not interchangeable, that is, one was the interviewer and the other the applicant), we computed a Pearson correlation coefficient to assess independence of dyad members. Consistent with Kenny's model, if scores are correlated, they are said to be non-independent and the dyad is the unit of analysis. If dyad scores do not significantly correlate, they are said to be independent and the unit of analysis can be the person. Analyses showed that interviewers' and applicants' implicit stereotyping scores were independent, $r = .23, p = .23$. The same was found for interviewers' and applicants' explicit stereotyping scores, $r = -.14, p = .44$. As such, we used the person as the level of analysis in subsequent analyses.

Using Hayes' PROCESS macro (Preacher & Hayes, 2008), and specifically a recent adaptation of this macro (Hayes, 2014), we tested a moderated serial mediation model (see Figure 1) in which implicit stereotypes were the predictor variable (X), explicit stereotypes (W1) and dummy-coded role (interviewer vs. applicant; W2) were both moderator variables, interviewer evaluations (M1) and applicant self-evaluations (M2) were serial mediators, and external performance evaluations were the outcome variable (Y). Age was included in the model as a control variable. Implicit stereotypes and explicit stereotypes were mean centered before computing interaction terms.

We ran this model twice, one time with interviewer role coded as 0 and applicant role coded as 1 (Model 1, Table 2), and a second time with reversed dummy codes (Model 2, Table 3). By reversing dummy codes, we were able to (a) investigate whether the serial indirect effect from implicit stereotypes to external performance evaluations was significant for interviewers' implicit stereotypes (Model 1), but not for applicant's implicit stereotypes (Model 2), as predicted and (b) probe any two-way and three-way interactions within the model, as is commonly done in hierarchical linear regression models, consistent with Aiken and West's (1991) method. For example, the main effect coefficient of a variable that significantly interacts with the dummy coded variable is the simple slope coefficient of that variable for the group coded 0.

Given that analyses were done at the individual level, and role (applicant vs. interviewer) was dummy coded, the overall sample size for the model was 60. The power of detecting a relationship was .92, as computed with the G*Power software. In terms of the overall model, the ratio of participants per parameters in the model is 10:1 (60 participants for a model with 6 parameters), which is considered acceptable.

Hypothesis Testing

Hypothesis 1. Our first hypothesis predicted that male interviewers' implicit stereotypes would indirectly predict external performance evaluations of female applicants, such that the effect of interviewers' implicit stereotypes on applicant's performance would be mediated, in a serial fashion, by interviewers' evaluations of the applicant and applicants' self-evaluations. Starting with interviewers' evaluations, our findings showed an interaction between implicit stereotypes and role predicting interviewers' evaluation of applicant, $b = 21.15, p = .05$. As can be seen in Model 1, interviewers' higher implicit stereotypes were associated with lower interviewer evaluations of female applicants, $b = -28.98, p = .003$. In other words, the more male interviewers associated women with incompetence (and men with

competence), the less hireable they evaluated the female applicant. Using Model 2 in which role dummy codes were reversed, we found that applicants' implicit stereotypes did not significantly predict interviewers' evaluations of applicants, $b = -7.82, p = .14$.

Next, findings revealed that male interviewers' evaluations of applicants predicted female applicants' self-evaluations, $b = .47, p = .001$, such that the less hireable the male interviewer evaluated the female applicant, the less hireable she evaluated herself following the interview. In turn, female applicants' self-evaluations predicted external evaluations of performance, $b = .25, p = .009$. In other words, the less hireable the applicant evaluated herself, the less hireable she was evaluated by external coders.

Finally, to establish mediation, we looked at the indirect effect from implicit stereotypes to externally rated performance evaluations through interviewers' evaluations and applicants' self-evaluations. This indirect effect was significant for interviewers' implicit stereotypes, as suggested by the confidence interval of the indirect effect, $b = -3.43, SE = 1.86, 95\% CI [-7.73, -0.36]$ with 10,000 bootstrap samples. However, the indirect effect was not significant for applicants' implicit stereotypes, $b = -0.93, SE = 1.04, 95\% CI [-3.54, 0.65]$ with 10,000 bootstrap samples. In other words, interviewers' (but not applicants') implicit stereotypes indirectly predicted external evaluators' performance evaluations of female applicants. These findings provide support for Hypothesis 1.

Hypothesis 2. Our second hypothesis predicted that interviewers' implicit and explicit gender stereotypes would interact to predict female applicants' job interview performance, such that female applicants would show the lowest performance when interacting with male interviewers who hold high implicit and low explicit gender stereotypes. We found a three-way interaction among implicit stereotypes, explicit stereotypes, and role in predicting external evaluations, $b = -0.63, p = .02$. When we broke down this interaction we found that interviewers' implicit and explicit stereotypes interacted

significantly to predict external evaluations for interviewers, $b = 0.59, p = .02$. However, applicants' implicit and explicit stereotypes did not interact, $b = -0.04, p = .68$.

Next, we broke down the significant interaction between interviewers' implicit and explicit gender stereotypes in predicting female applicants' externally rated performance. To understand this significant interaction, we used Aiken and West's (1991) method of testing simple slopes for interviewers' implicit gender stereotypes at high and low levels of interviewers' explicit gender stereotypes. High and low levels were computed at one standard deviation above and below the mean of interviewers' explicit gender stereotypes. Results are summarized in Figure 2. For interviewers displaying high explicit gender stereotypes, their implicit gender stereotypes did not predict female applicants' externally rated performance, $b^* = .40, p = .13$. However, when interviewers were low in explicit gender stereotypes, higher implicit gender stereotypes predicted lower performance of female applicants, as evaluated by external coders, $b^* = -.59, p = .015$. Thus, female applicants are most vulnerable to interviewers' negative implicit stereotypes when they are coupled with low (weak) explicit gender stereotypes. These findings provide support for Hypothesis 2.

Hypothesis 3. Our final hypothesis predicted that female applicants' own implicit gender stereotypes would directly predict their externally rated performance, such that the more female job applicants associate women with incompetence, the less competently they would perform in the job interview. We found an interaction between implicit stereotypes and role in predicting external evaluations of applicant, $b = -15.12, p = .04$. As suggested by Model 1, interviewers' implicit stereotypes were not directly associated with external evaluations, $b = 4.60, p = .50$. (Instead, as predicted in Hypothesis 1, they indirectly predicted external evaluations.) However, as suggested by Model 2, applicants' higher implicit stereotypes were directly associated with lower external evaluations. $b = -10.52, p = .005$. In other words, applicants' implicit stereotypes directly predicted how they were evaluated by

external coders who evaluated their performance based on videotapes. These findings provide support for Hypothesis 3.

Discussion

The goal of the present study was to investigate how implicit and explicit stereotypes held by both male interviewers and female applicants predicted women's performance outcomes in job interviews. Several important findings emerged. First, male interviewers' and female applicants' implicit gender stereotypes both predicted the performance of female applicants. However they did so in different ways. Whereas interviewers' implicit stereotypes predicted their own evaluations of female applicants' performance, applicants' implicit stereotypes predicted externally evaluated performance (the impression the applicant made to external viewers). However, interviewers' implicit gender stereotypes also indirectly predicted externally evaluated performance, through interviewer evaluations and applicant self-evaluations, showing evidence of a self-fulfilling prophecy. This finding suggests that for interviewers' implicit gender stereotypes to predict female applicants' performance, there needs to be a social interaction in which interviewers' stereotypical expectations are communicated to the female applicants, hence the indirect process. The present study offers a more precise picture of this process: Interviewers' implicit associations between gender and incompetence, although not fully conscious, guided interviewers' evaluations of the applicants' performance. The female applicants, in turn, seemed to assimilate the impression of the interviewers in evaluating their own performance. These self-evaluations predicted how hireable applicants were rated by external raters.

Applicants' own implicit gender stereotypes directly influenced how applicants performed. Women who had implicit negative associations of women behaved in a way that made them seem less hireable. This finding adds to the work on stereotype threat, which shows that being aware of a negative stereotype affects performance (Spencer, Steele, &

Quinn, 1999; Steele & Aronson, 2000). We extend these findings by showing that women who implicitly endorse negative stereotypes about women tend to perform less well.

However, it is interesting to note that female applicants' implicit gender stereotypes did not significantly predict their self-evaluations of performance. Instead, women's implicit stereotypes directly and exclusively predicted women's behavior as judged by external coders. This finding suggests that women may not be aware of either their own negative implicit stereotypes or the way in which these stereotypes influence their behavior.

Second, it is important to note that explicit stereotypes did not generally predict performance outcomes independently. However, when looking at the interplay between interviewers' implicit and explicit stereotypes, we found that they interacted to predict external evaluations of performance, such that having an interviewer who was at the same time high in implicit and low in explicit stereotypes predicted the lowest performance of female applicants. This is consistent with work on interracial interactions (Dovidio, 2001; Penner et al., 2010), which shows that Whites' combination of high implicit/low explicit racial biases leads to the most negative consequences in interracial interactions. Dovidio and colleagues (Dovidio & Gaertner, 2000; Penner et al., 2010) called these individuals "aversive racists" and speculated that their incongruent verbal and nonverbal behaviors may account for the disruptive effects in the interaction. Similarly, aversive sexists may display positive verbal behaviors (stemming from their explicit stereotypes) and negative nonverbal behaviors (stemming from their implicit stereotypes). This inconsistency may be especially taxing for female job applicants who are trying to decipher their interviewers' signals, and resolving this inconsistency may have affected women's interview performance.

Third, the current findings are in contrast with studies that found that implicit measures (as assessed by an IAT) were worse (Karpinski & Hilton, 2001) or no better predictors (Oswald, Mitchell, Blanton, Jaccard, & Tetlock, 2013) of behavior compared to

explicit measures. One reason why our findings may differ is that we used a different task to measure implicit associations (a sequential priming task), which may be better suited for measuring predictor variables. On a more theoretical level, we suggest that considering both implicit and explicit biases in interaction would best predict outcomes, especially in actual social interactions in which verbal and nonverbal signals are freely expressed.

Theoretical Implications

Our findings have important implications for the literature on the self-fulfilling prophecy because they pinpoint the possible sources of self-fulfilling prophecies in mixed-gender interactions. First, we show that one source is interviewers' implicit gender stereotypes, which can confirm themselves in the performance of female applicants. Moreover, interviewers' negative implicit stereotypes are especially harmful when coupled with positive explicit stereotypes. Second, we show that female applicants' implicit gender stereotypes can also lead to self-fulfilling prophecy effects: Women who believe that women are not competent managers tend to perform less well, thus confirming their self-stereotypes. Overall, these findings add to the self-fulfilling prophecy literature by showing that beyond perceived interviewer biases (von Baeyer et al., 1981), actual interviewer and applicant biases can predict how applicants perform.

This latter finding is also important for the stereotype threat literature, which shows that the awareness of a negative stereotype can undermine performance on tasks for which the stereotype is relevant. We further show that implicitly endorsing negative stereotypes, for both interaction partners, can lead to performance decrements. This finding also adds to the stereotype threat literature, which usually does not study the effects of stereotypes in actual social interactions (i.e., stereotype threat effects can be obtained in the absence of an actual interaction, for example by making gender salient in a testing situation; Steele & Aronson, 1995, Study 4).

Finally, the current findings are consistent with Heilman's (2001) work on how gender stereotypes influence evaluations of women in managerial roles, due to their descriptive role (stereotypes dictate what women are like) and their prescriptive role (stereotypes dictate how women should be). For example, consistent with Heilman's lack-of-fit model, we found that gender stereotypes led interviewers to devalue female applicants' performance and created negative performance expectations which confirmed themselves in the performance of female applicants. Our findings add to Heilman's work in two ways. First, they suggest that the stereotypes of evaluators (in our case interviewers) can predict not only their evaluations of women, but also women's own gender stereotypes. Second, our findings suggest that distinguishing between implicit and explicit gender stereotypes is important.

Practice Implications

Pinpointing the source of self-fulfilling prophecies is important for practical purposes, as it suggests specific avenues for reducing this effect. One option is to train interviewers to decrease implicit stereotypes using empirically tested training techniques such as the Just Say No Training (Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000) or Situational Attribution Training (Stewart, Latu, Kawakami, & Myers, 2010). These training techniques are especially designed to reduce implicit negative stereotypes, either by negating stereotypes or by learning to make situational attributions for negative behaviors. Although these training techniques have been designed for reducing different stereotypes (racial, elderly, skinheads), they can likely be adapted to reduce implicit gender stereotypes.

Having established mediating factors of the effect, we can also suggest several points in the process at which we can intervene. For example, given that interviewers' evaluations mediate the effect, interviewers can be trained to assess performance more objectively. Also, a more empowering strategy for female applicants would be to separate their self-evaluations from indirect performance feedback coming from the interviewer. Given that female

applicants' own implicit stereotypes also predicted their performance impressions, another empowering strategy may be to train applicants themselves to increase implicit associations between women and managerial competence through repeated exposures to women-competence associations.

We showed negative effects of implicit gender stereotypes in a semi-structured interview. According to Heilman (2001), ambiguity in evaluation criteria and a lack of structure in the evaluation process—both of which are present in unstructured or semi-structured interview situations—are factors which can facilitate the negative effects of stereotypes in evaluating women. Because of this, we suggest that it may be possible to reduce the negative effects of implicit gender stereotypes when job interviews are structured and evaluation criteria are highly specified. This recommendation would be consistent with research showing that bias against individuals from different disadvantaged groups (overweight individuals, individuals with physical disabilities, pregnant women) was reduced in structured job interviews (Bragger, Kutcher, Morgan, & Firth, 2002; Brecher, Bragger, & Kutcher, 2006; Kutcher & Bragger, 2004).

Limitations and Future Directions

A few limitations of our study should be noted. First, because we measured rather than manipulated interviewers' implicit stereotypes, we were unable to establish direct causation. However, this design may also be construed as a strength. Jussim (1990) criticized the self-fulfilling prophecy literature for using experimenter-manipulated expectations that might be false (not corresponding to the actual expectations of participants). As Jussim noted, we have "little information about the extent to which naturally occurring expectations create social reality and contribute to social problems" (p. 30). We bypass this issue by showing that naturally occurring stereotypes of both interaction partners can predict outcomes in actual social interactions.

Also, despite our attempts to employ a cover story that separated the two phases of the study, the measurement of explicit sexism in the first phase may have made gender more salient in the job interview phase. This salience, however, would have occurred for all participants and can therefore not be responsible for explaining the systematic relation between interviewers' implicit stereotypes and the performance of applicants. However, in future studies researchers might consider avoiding the measurement of explicit sexism before the targeted social interactions.

We are still unclear about the actual verbal and/or nonverbal mechanism through which interviewers' evaluations are communicated to female applicants. Within the racial interaction literature, White participants' implicit racial stereotypes were expressed through nonverbal behaviors suggesting lack of friendliness (Dovidio et al., 2002). However, unlike interracial interactions, men and women are highly interdependent and interpersonally close; they form friendships, date, marry, live together, and raise families. In fact, the intergroup anxiety that is characteristic of interracial interactions (Finchilescu, 2010; Plant & Devine, 2003) is less pronounced in mixed-gender interactions (Blair, Park, & Bachelor, 2003). As such, nonverbal friendliness may not be diagnostic in mixed-gender interactions. Instead, it is more likely that implicit gender stereotypes would "leak" through nonverbal behaviors related to dominance. However, in the context of the present study, it was difficult to study this hypothesis given that both interaction partners were naïve participants. Given the dynamic nature of the interaction, we cannot determine whether a certain verbal or nonverbal behavior of a male interviewer is determined by his implicit gender biases or is a reaction to the applicant's behavior. Consistent with Dovidio and colleagues (2002), future studies should employ confederates as applicants, who are trained to respond in a scripted way. Thus, variations in interviewers' verbal/nonverbal behaviors would likely be attributable to their implicit gender stereotypes and not to the applicants' performance.

In addition, performance evaluations were measured only after the interview. Because of this procedure, it is possible, for example, that applicants' self-evaluations were affected by perceived poor performance, rather than self-evaluations leading to performance being perceived poorly. In future studies, it would be helpful to examine evaluations at different points during the interview in order to get a more precise view of the social interaction as it unfolds.

Finally, in the current study we included only women as targets of bias. From an applied perspective, this is important given that women are more likely to face disadvantages during the hiring process compared to men and, as such, it is important to understand how gender biases contribute to these disadvantages. However, from a theoretical point of view, it will be interesting in future research to look at the effects of interviewers' gender biases on the performance of male applicants. We can speculate that men, given that they benefit from the advantage of being in the powerful group, do not experience stereotype threat, but instead are advantaged by the stereotype—a phenomenon known as *stereotype lift* (Walton & Cohen, 2003). Not being in a threatening situation, men may be less likely to be negatively influenced by their interviewers' gender stereotypes. Another reason why we speculate men may not be victims of their interviewers' implicit stereotypes is that men are less sensitive to their partners' expectations compared to women and thus less likely to show self-fulfilling prophecy effects (Christensen & Rosenthal, 1982; Lippa, 2005). However, men's personal implicit gender stereotypes may still predict their performance in the job interview, such that believing that men are less likely to be associated with leadership may directly affect men's self-efficacy, or even their demeanor, during the task.

Conclusions

There has been progress in terms of gender equality in the last few decades. Moreover, gender stereotypes, particularly explicit ones, have become less negative than in

the past. However, this does not mean that gender biases no longer exist. As Gaertner and Dovidio (1986, p. 85) put it, prejudice is “like a virus that mutates into new forms.” One such new form—implicit gender stereotypes—seems to play a vital role in predicting women’s outcomes in evaluative situations such as job interviews. Whether it is the implicit gender stereotypes of male interviewers or female applicants, these relatively unconscious and hard-to-control associations can have negative consequences for women’s outcomes in the workplace.

References

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA: Sage Publications, Inc.
- Blair, I. V., Park, B., & Bachelor, J. (2003). Understanding intergroup anxiety: Are some people more anxious than others? *Group Processes & Intergroup Relations*, 6(2), 151-169. doi: 10.1177/1368430203006002002
- Bommer, W. H., Johnson, J., Rich, G. A., Podsakoff, P. M., & MacKenzie, S. B. (1995). On the interchangeability of objective and subjective measures of employee performance: A meta-analysis. *Personnel Psychology*, 48(3), 587-605. doi: 10.1111/j.1744-6570.1995.tb01772.x
- Bragger, J. D., Kutcher, E., Morgan, J., & Firth, P. (2002). The effects of the structured interview on reducing biases against pregnant job applicants. *Sex Roles*, 46(7-8), 215-226. doi: 10.1023/a:1019967231059
- Brecher, E., Bragger, J., & Kutcher, E. (2006). The structured interview: reducing biases toward job applicants with physical disabilities. *Employee Responsibilities and Rights Journal*, 18(3), 155-170.
- Catalyst. (2011). *U.S. Women in Business*. Retrieved from <http://www.catalyst.org/publication/132/us-women-in-business>
- Christensen, D., & Rosenthal, R. (1982). Gender and nonverbal decoding skill as determinants of interpersonal expectancy effects. *Journal of Personality and Social Psychology*, 42(1), 75-87. doi: 10.1037/0022-3514.42.1.75
- Cook, W. L., & Kenny, D. A. (2005). The Actor-Partner Interdependence Model: A model of bidirectional effects in developmental studies. *International Journal of Behavioral Development*, 29(2), 101-109. doi: 10.1080/01650250444000405

- Dovidio, J. F. (2001). On the nature of contemporary prejudice: The third wave. *Journal of Social Issues*, 57(4), 829-849. doi: 10.1111/0022-4537.00244
- Dovidio, J. F., & Gaertner, S. L. (2000). Aversive racism and selection decisions: 1989 and 1999. *Psychological Science*, 11(4), 315-319. doi: 10.1111/1467-9280.00262
- Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology*, 82(1), 62-68. doi: 10.1037/0022-3514.82.1.62
- Duehr, E. E., & Bono, J. E. (2006). Men, women, and managers: Are stereotypes finally changing? *Personnel Psychology*, 59(4), 815-846. doi: 10.1111/j.1744-6570.2006.00055.x
- Finchilescu, G. (2010). Intergroup anxiety in interracial interaction: The role of prejudice and metastereotypes. *Journal of Social Issues*, 66(2), 334-351. doi: 10.1111/j.1540-4560.2010.01648.x
- Freund, P. A., & Kasten, N. (2012). How smart do you think you are? A meta-analysis on the validity of self-estimates of cognitive ability. *Psychological Bulletin*, 138(2), 296-321. doi: 10.1037/a0026556
- Gaertner, S. L., & Dovidio, J. F. (1986). The aversive form of racism. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism*. (pp. 61-89). San Diego, CA: Academic Press.
- Glick, P., & Fiske, S. T. (2001). An ambivalent alliance: Hostile and benevolent sexism as complementary justifications for gender inequality. *American Psychologist*, 56, 109 - 118.
- Hayes, A. F. (2014). *An index and test of linera moderated mediation*. Manuscript submitted for publication.

- Heilman, M. E. (2001). Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues*, 57(4), 657-674. doi: 10.1111/0022-4537.00234
- Heilman, M. E., Block, C. J., & Martell, R. F. (1995). Sex stereotypes: Do they influence perceptions of managers? *Journal of Social Behavior & Personality*, 10(6), 237-252.
- Howard, J. L., & Ferris, G. R. (1996). The employment interview context: Social and situational influences on interviewer decisions. *Journal of Applied Social Psychology*, 26(2), 112-136. doi: 10.1111/j.1559-1816.1996.tb01841.x
- John, O. P., & Robins, R. W. (1994). Accuracy and bias in self-perception: Individual differences in self-enhancement and the role of narcissism. *Journal of Personality and Social Psychology*, 66(1), 206-219. doi: 10.1037/0022-3514.66.1.206
- Jussim, L. (1990). Social reality and social problems: The role of expectancies. *Journal of Social Issues*, 46(2), 9-34. doi: 10.1111/j.1540-4560.1990.tb01921.x
- Karpinski, A., & Hilton, J. L. (2001). Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology*, 81(5), 774-788. doi: 10.1037/0022-3514.81.5.774
- Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): Effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology*, 78(5), 871-888. doi: 10.1037/0022-3514.78.5.871
- Koenig, A. M., Eagly, A. H., Mitchell, A. A., & Ristikari, T. (2011). Are leader stereotypes masculine? A meta-analysis of three research paradigms. *Psychological Bulletin*, 137(4), 616-642. doi: 10.1037/a0023557
- Kutcher, E. J., & Bragger, J. D. (2004). Selection interviews of overweight job applicants: Can structure reduce the bias? *Journal of Applied Social Psychology*, 34(10), 1993-2022. doi: 10.1111/j.1559-1816.2004.tb02688.x

Latu, I. M., Stewart, T. L., Myers, A. C., Lisco, C. G., Estes, S. B., & Donahue, D. K. (2011).

What we “say” and what we “think” about female managers: Explicit versus implicit associations of women with success. *Psychology of Women Quarterly*, 35(2), 252-266. doi: 10.1177/0361684310383811

Lippa, R. A. (2005). *Gender, nature, and nurture* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Nosek, B. A., Banaji, M., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice*, 6(1), 101-115. doi: 10.1037/1089-2699.6.1.101

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies. *Journal of Personality and Social Psychology*, 105(2), 171-192. doi: 10.1037/a0032734
10.1037/a0032734.supp (Supplemental)

Penner, L. A., Dovidio, J. F., West, T. V., Gaertner, S. L., Albrecht, T. L., Dailey, R. K., ... Markova, T. (2010). Aversive racism and medical interactions with Black patients: A field study. *Journal of Experimental Social Psychology*, 46(2), 436-440. doi: 10.1016/j.jesp.2009.11.004

Plant, E. A., & Devine, P. G. (2003). The antecedents and implications of interracial anxiety. *Personality and Social Psychology Bulletin*, 29(6), 790-801. doi: 10.1177/0146167203029006011

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40(3), 879-891. doi: 10.3758/brm.40.3.879

Prediger, D. J. (1999). Basic structure of work-relevant abilities. *Journal of Counseling Psychology*, 46(2), 173-184. doi: 10.1037/0022-0167.46.2.173

- Ridge, R. D., & Reber, J. S. (2002). 'I think she's attracted to me': The effect of men's beliefs on women's behavior in a job interview scenario. *Basic and Applied Social Psychology, 24*(1), 1-14. doi: 10.1207/153248302753439029
- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. New York, NY: Holt, Rinehart & Winston.
- Rudman, L. A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *Journal of Social Issues, 57*(4), 743-762. doi: 10.1111/0022-4537.00239
- Rudman, L. A., & Kilianski, S. E. (2000). Implicit and explicit attitudes toward female authority. *Personality and Social Psychology Bulletin, 26*(11), 1315-1328. doi: 10.1177/0146167200263001
- Rudman, L. A., & Phelan, J. E. (2010). The effect of priming gender roles on women's implicit gender beliefs and career aspirations. *Social Psychology, 41*(3), 192-202. doi: 10.1027/1864-9335/a000027
- Schein, V. E., Mueller, R., & Jacobson, C. (1989). The relationship between sex role stereotypes and requisite management characteristics among college students. *Sex Roles, 20*(1-2), 103-110. doi: 10.1007/bf00288030
- Schein, V. E., Mueller, R., Lituchy, T., & Liu, J. (1996). Think manager—think male: A global phenomenon? *Journal of Organizational Behavior, 17*(1), 33-41. doi: 10.1002/(sici)1099-1379(199601)17:1<33::aid-job778>3.0.co;2-f
- Schmid Mast, M. (2004). Men are hierarchical, women are egalitarian: An implicit gender stereotype. *Swiss Journal of Psychology/Schweizerische Zeitschrift für Psychologie/Revue Suisse de Psychologie, 63*(2), 107-111. doi: 10.1024/1421-0185.63.2.107

- Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science, 10*(1), 80-83. doi: 10.1111/1467-9280.00111
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*(1), 4-28. doi: 10.1006/jesp.1998.1373
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*(5), 797-811. doi: 10.1037/0022-3514.69.5.797
- Steele, C. M., & Aronson, J. (2000). Stereotype threat and the intellectual test performance of African Americans. In C. Stangor (Ed.), *Stereotypes and prejudice: Essential readings* (pp. 369-389). New York, NY: Psychology Press.
- Stewart, T. L., Latu, I. M., Kawakami, K., & Myers, A. C. (2010). Consider the situation: Reducing automatic stereotyping through situational attribution training. *Journal of Experimental Social Psychology, 46*(1), 221-225. doi: 10.1016/j.jesp.2009.09.004
- Stoker, J. I., Van der Velde, M., & Lammers, J. (2012). Factors relating to managerial stereotypes: The role of gender of the employee and the manager and management gender ratio. *Journal of Business and Psychology, 27*(1), 31-42. doi: 10.1007/s10869-011-9210-0
- von Baeyer, C. L., Sherk, D. L., & Zanna, M. P. (1981). Impression management in the job interview: When the female applicant meets the male (chauvinist) interviewer. *Personality and Social Psychology Bulletin, 7*(1), 45-51. doi: 10.1177/014616728171008
- Walton, G. M., & Cohen, G. L. (2003). Stereotype lift. *Journal of Experimental Social Psychology, 39*(5), 456-467. doi: 10.1016/s0022-1031(03)00019-2

Williams, M. J., Paluck, E. L., & Spencer-Rodgers, J. (2010). The masculinity of money:

Automatic stereotypes predict gender differences in estimated salaries. *Psychology of*

Women Quarterly, 34(1), 7-20. doi: 10.1111/j.1471-6402.2009.01537.x

Word, C. O., Zanna, M. P., & Cooper, J. (1974). The nonverbal mediation of self-fulfilling

prophecies in interracial interaction. *Journal of Experimental Social Psychology*,

10(2), 109-120. doi: 10.1016/0022-1031(74)90059-6

Footnotes

¹Although our research was guided by models in which mixed-group interactions are the norm (e.g., White/Black interactions; Dovidio et al., 2002), we also tested several female/female (same-group) dyads. Preliminary analyses of 16 female interviewer/female applicant dyads showed no significant effects. Although power for these analyses was low due to the small sample size (e.g., .61 for interviewer implicit stereotypes), effect sizes were also consistently low. For example, *Cohen's f*² values for interviewers' implicit stereotypes were .02 for predicting interviewers' evaluations, .0005 for predicting applicants' self-evaluations, and .15 for predicting external evaluation. For applicants' implicit stereotypes, *Cohen's f*² values were .15 for predicting interviewers' evaluations, .01 for predicting applicants' self-evaluations, and .07 for predicting external evaluation. To investigate whether an increase in sample size would lead to significant effects, we also performed bootstrapping on all the analyses and found no significant results (all *ps* > .05). Given these preliminary findings, as well as the models of mixed-group interactions on which we based our research, we focused exclusively on the mixed-gender (male interviewer/female applicant) dyads.

²The term self-fulfilling prophecy is often used interchangeably with the term behavioral confirmation (e.g., Lippa, 2005). Although we describe studies that have used both terms, we prefer using the more general term of self-fulfilling prophecy and avoid the term behavioral confirmation because our outcome is not necessarily behavior per se, but rather the impressions that female applicants make in terms of performance.

³Details about the Hostile and Benevolent Sexism scales and results are presented in the Supplementary Material section.

Table 1

Descriptive Statistics and Correlations of Study 1 Variables for Male and Female Participants

Variables	Male Interviewers (<i>n</i> = 30)		Female Applicants (<i>n</i> = 30)		Correlations					
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	1	2	3	4	5	6
1. Implicit Stereotypes	-0.001	0.02	0.003	0.04	-	-.07	-.52**	-.34	-.17	.009
2. Explicit Stereotypes	14.83	24.05	29.87	29.62	-.03	-	.19	.28	.18	-.27
3. Interviewer Evaluation	5.14	1.29	5.14	1.29	-.28	-.005	-	.51**	.21	.18
4. Applicant Self-Evaluation	4.29	1.35	4.29	1.35	-.29	-.08	.51**	-	.36	.26
5. External Evaluation	3.39	0.95	3.39	0.95	-.53**	-.05	.21	.36	-	-.31
6. Age	25.80	3.90	23.80	2.71	-.17	-.03	.05	.22	-.07	-

Note. Intercorrelations of study variables for male interviewers are presented above the diagonal and for female applicants, below the diagonal.

p* < .05. *p* < .01.

Table 2

Moderated serial mediation, Model 1

	Male Interviewer Evaluation of Applicant (<i>M1</i>)		Female Applicant Self-evaluation (<i>M2</i>)		External Evaluation of Applicant (<i>Y</i>)	
	Coeff.	<i>p</i>	Coeff.	<i>p</i>	Coeff.	<i>p</i>
Implicit Stereotypes (<i>X</i>)	-28.98	.003	-3.86	.70	4.60	.50
Explicit Stereotypes (<i>W1</i>)	0.003	.64	0.003	.55	0.002	.69
Role: int. = 0; appl. = 1 (<i>W2</i>)	0.14	.67	0.10	.76	-0.22	.32
Interviewer Evaluation (<i>M1</i>)			.47	.001	0.01	.89
Applicant Self-evaluation (<i>M2</i>)					0.25	.009
X x W1	-0.30	.40	0.14	.69	0.59	.02
X x W2	21.15	.05	-0.54	.96	-15.12	.04
X x W1 x W2	0.11	.78	-0.10	.80	-0.63	.02
Age (<i>U</i>)	0.05	.28	0.07	.14	-0.10	.004
Constant	3.75	.004	0.04	.97	4.79	.00
	$R^2 = .22$		$R^2 = .31$		$R^2 = .40$	
	$F(7, 52) = 2.16, p = .05$		$F(8, 51) = 2.94, p = .009$		$F(9, 50) = 3.79, p = .001$	

Table 3

Moderated Serial Mediation, Model 2

	Male Interviewer Evaluation of Applicant (<i>M1</i>)		Female Applicant Self-evaluation (<i>M2</i>)		M1 External Evaluation of Applicant (<i>Y</i>)	
	Coeff.	<i>p</i>	Coeff.	<i>p</i>	Coeff.	<i>p</i>
Implicit Stereotypes (<i>X</i>)	-7.82	.14	-4.41	.41	-10.52	.005
Explicit Stereotypes (<i>W1</i>)	0.003	.64	0.004	.55	0.002	.69
Role: int. = 1; appl. = 0 (<i>W2</i>)	-0.14	.67	-0.10	.76	0.22	.32
Interviewer Evaluation (<i>M1</i>)			0.47	.001	0.01	.89
Applicant Self-evaluation (<i>M2</i>)					0.25	.009
X x W1	-0.19	.25	0.04	.81	-0.04	.68
X x W2	-21.15	.05	0.54	.96	15.12	.04
X x W1 x W2	-0.11	.78	0.10	.80	0.63	.02
Age (<i>U</i>)	0.05	.28	0.07	.14	-0.10	.004
Constant	3.89	.002	-0.14	.91	4.57	.00
	$R^2 = .22$		$R^2 = .31$		$R^2 = .40$	
	$F(7, 52) = 2.16, p = .05$		$F(8, 51) = 2.94, p = .009$		$F(9, 50) = 3.79, p = .001$	

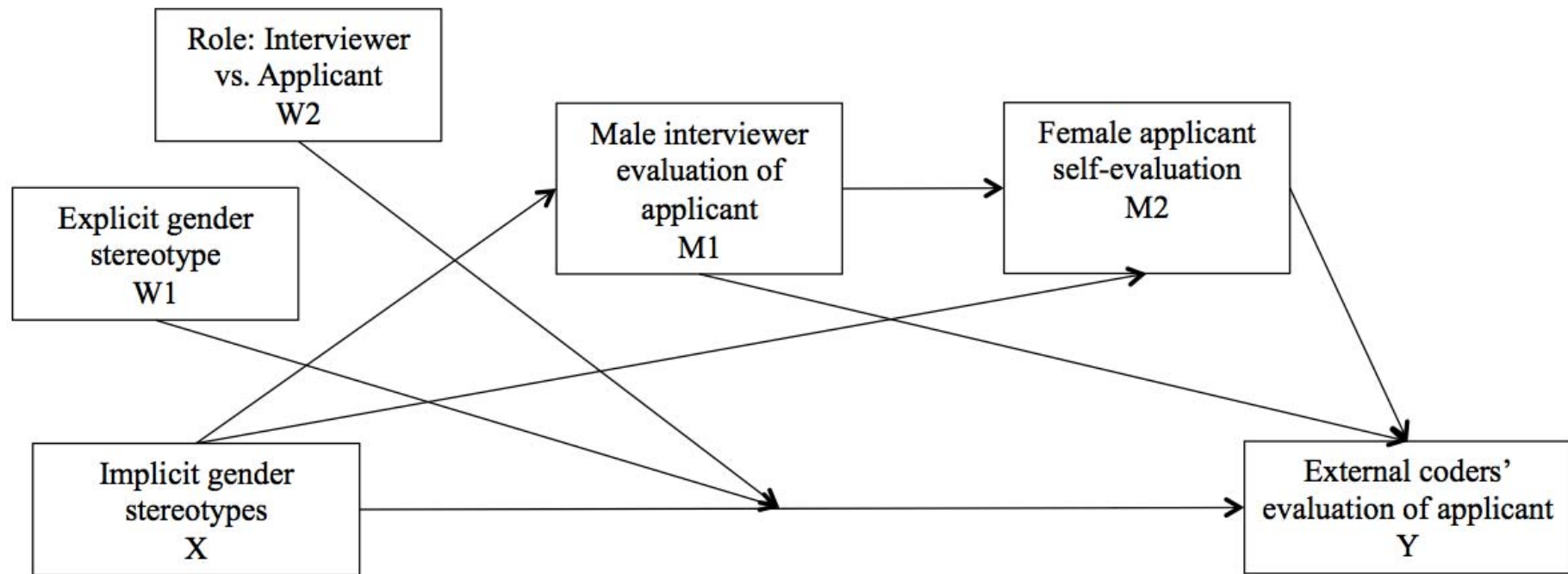


Figure 1. Moderated Serial Moderation Conceptual Model

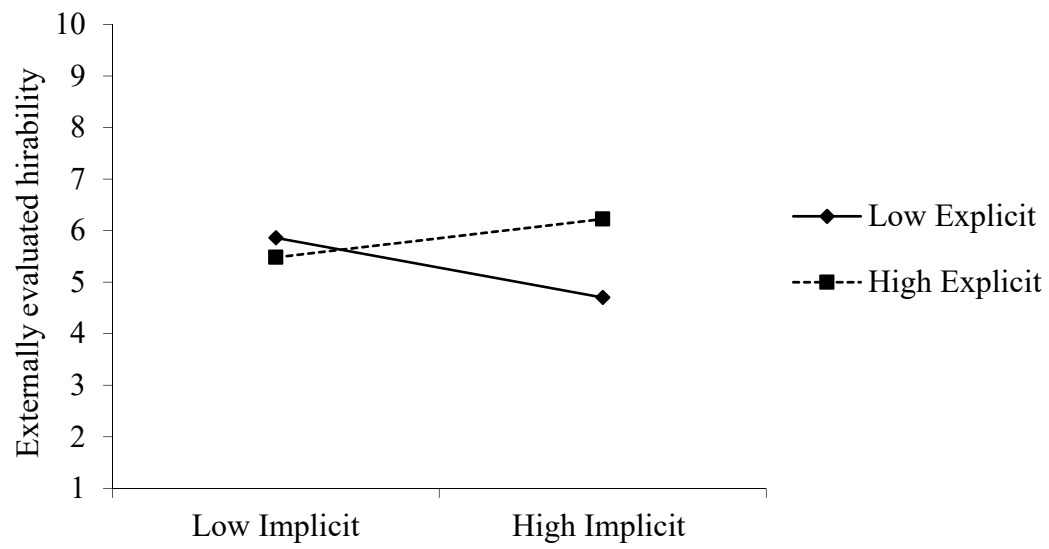


Figure 2. Interaction between interviewer implicit and explicit gender stereotypes in predicting externally rated hirability